图书情报工作
第64卷 第10期 2020年5月

科学数据相关性标准使用特征*

■ 张贵兰[1] 周国民[2] 潘尧[2] 刘建平[2] 王健[2]
[1] 中国科学技术信息研究所 北京 100038 [2] 中国农业科学院农业信息研究所 北京 100081

摘 要：[目的/意义]研究以科学数据为信息载体，探究用户在进行数据检索时，影响其进行相关性判断的标准的使用特征，从而进一步探究科学数据用户的信息行为，为完善数据检索平台和加强数据共享与服务提供理论支持。[方法/过程]通过调查问卷、访谈与层次分析法相结合的方法开展研究，用户为高校和科研单位在读研究生。[结果/结论]分析结果显示相关性标准呈现4个特征：①相关性标准的使用情况符合著名的长尾定律；②主题性在相关性判断过程中具有锚定功能；③相关性标准在相关性判断中具有双向性；④相关性标准使用权重相对稳定。研究准确把握了科学数据用户的检索策略，明晰了其检索行为，为改进科学数据共享系统提供了具有可操作性的建议。

关键词：科学数据 相关性标准 相关性标准使用特征 用户行为
分类号：G251
DOI：10.13266/j.issn.0252-3116.2020.10.007

1 引言

相关性标准的研究起源于上世纪60年代，二战之后随着计算机的发展迎来了信息爆炸。为了提高信息检索效率，学者开始关注相关性标准（C. A. Cuadra等[1]和A. M. Ree等[2]）。随后经过将近20年的低谷期，1990年后随着互联网时代的到来，迎来了相关性标准研究的复苏和繁荣发展。主要代表学者是C. Barry、L. Schamber等，他们的对比研究发现存在跨越不同情境而存在的共性标准[3]。A. R. Tarloy等[4]、Y. Calvin等[5]学者发现在不同情境下，相关性标准的集合差异不大，但是相关性标准的使用变化较大。由此可见相关性标准使用特征直接影响着用户的相关性判断结果。

随着数据时代的到来，科学数据已经成为科技研究领域重要的资源，科学数据共享不仅提高了科研人员的工作效率，更避免了国家资源的重复浪费。如今大批数据共享平台建立，但是其检索结果却不尽人意。如何高效地从中检索与用户需求相关的数据成为亟待解决的问题。本文将从情境相关性角度入手，开展科学数据相关性标准使用特征的研究，力图从用户的角度为数据共享平台的改进提供建议。

本文基于科学数据使用特征开展研究，研究的问题主要有以下3个：①科学数据用户相关性判断中标准的使用频率、重要性和顺序；②科学数据用户在进行相关性判断过程中，各个相关性标准所起的作用；③在不同的数据相关性判断情境中，相关性标准使用的差异性。

2 文献综述

相关性标准的研究发展主要呈现两大特点：

（1）相关性标准研究的信息载体都是迎合时代的发展，以当代主流的信息载体为研究对象。20世纪90年代，C. Barry[6]，L. Schamber等[7]，C. Cool等[8]，P. Wang等[9-10]学者针对文献相关性标准开展了深入研究。进入21世纪，相关性标准的研究维度由文献扩到互联网、多媒体（视频、音乐、图像）、社交平台等。从2010年之后，其研究维度又进一步拓宽到数据领域

* 本文系中国农业科学院农业信息研究所基本科研业务费重点项目"科学数据出版能力建设研究"（项目编号：JBYW-AII-2020-35）和国家社会科学基金项目"基于信息熵的学术期刊传播规律和评价方法研究"（项目编号：15BTQ059）研究成果之一。
作者简介：张贵兰（ORCID：0000-0002-9153-3579），助理研究员，博士；周国民（ORCID：0000-0001-9906-8061），研究员，博士；潘尧（ORCID：0000-0003-2317-7442），硕士研究生；刘建平（ORCID：0000-0002-1817-5373），博士研究生；王健（ORCID：0000-0003-4958-7669），研究员，博士，通讯作者，E-mail：wangjian01@caas.cn。
收稿日期：2019-11-06 修回日期：2020-02-07 本文起止页码：56-65 本文责任编辑：易飞

(地理数据、商务数据、科学数据)。

(2)相关性标准使用的研究总是在相关性标准集合研究之后。无论何种维度的相关性研究，总是先探究其集合，再探究其使用。20世纪90年代，相关性标准的研究重点在相关性标准是什么以及相关性标准有哪些的内容上。在相关性标准的研究基础上，后人开展了不同维度下、不同阶段、不同用户群体相关性标准的使用研究。

相关性标准的研究随着时代的发展不断丰富与完善。从最初的文字时代，到网络时代，再到现在的数据时代，相关性标准研究的信息载体从文献发展到网页、图像、视频、社交、音乐、电子商务以及科学数据等；研究情境从科研工作发展到日常生活、休闲娱乐、学生作业；研究对象从科研工作者发展到中学生、教师、网民、医生、记者等。

2.1 数据相关性标准研究

C. Barry 和 L. Schamber 研究对比发现存在跨域情境而存在的共性标准[3]，因此后续相关性标准的研究都是在此基础上进行不断修正、完善和补充的。网页的相关性标准增添了安全性、便利性；图像的相关性标准增添了分辨率、清晰性和方向性；多媒体的相关性标准增添了娱乐性。

随着数据时代的到来，针对数据的相关性标准也不断丰富。最先对数据开展研究的领域是电子商务。O. Alonso 等[11]研究了用户在亚马逊购买商品时进行电子商务相关性判断时所使用的相关性标准，并与非电子商务情境下使用的相关性标准频率进行了对比分析。研究结果显示电子商务情境中，会出现新的相关性标准(售后服务、价格等)，同时相关标准的使用情况也会有所不同。D. S. Sabbata 等[12－13]研究了用户在选择地理数据的时候使用的相关性标准，发现与传统的数据相关性标准有所不同，出现了地理方面特有的标准，比如可视化、空间性、方向性等。T. Reichenbacher 等[14]针对地理数据相关性开展了进一步研究，聚焦于移动手机地理判断，研究发现了五大类相关性标准——主题性、时空性、方向性、聚类和通信交换机，其中通信交换机是移动手机特有的判断标准。X. Wang 等[15]基于相关性判断和移动营销的基础，开展了移动商务信息开展相关性研究。研究采用任务调研的方式(vignette-based survey)，识别了5个相关性标准，分别是主题性、经济价值、位置、时间和可信度。

高飞[16]首次开展针对农业领域的科学数据的研究，借鉴传统学者 C. Barry 等和网页相关性研究方面 Balatsoukas 的实验设计方案，采用三角测量的方法，全方位采集用户的检索过程，包括事前访谈、情境实验、眼动仪记录、事后访谈、调查问卷等方式。最终识别了9个科学数据相关性标准，分别是主题性、质量、功能性、便利性、可理解性、可获得性、权威性、时效性、新颖性。

张贵兰等[17]在高飞研究的基础上对科学数据相关性标准进行了细化研究，系统梳理了相关性标准的研究文献，发现了相关性标准的两大特点：①存在跨域情境而存在的共性标准集合；②任何信息载体下都存在与自身特点相关的固有相关性标准。在此基础上通过情境实验和内容分析，识别了9个科学数据相关性标准。

2.2 相关性标准使用研究

J. Bateman[18－19]认为相关性标准是影响相关性判断的关键因素，并且会呈阶段性变化，他引入 Kuhlthau 提出的信息检索的6个认知阶段，尝试通过实证数据(调查问卷)来识别相关性标准并探究随着认知阶段的变化，相关性标准的变化情况。P. Vakkari 等[20]研究了相关性标准随任务情境变化的动态性。他借助真实的课堂环境，将课程分为开始、中间和结束三个阶段，并在每个阶段为11个学生布置检索任务，通过访谈与出声思考的方式获取数据。分析结果显示，用户无论是判断参考资料还是文献时使用的相关性标准均与任务情境有密切关系。K. L. Maglaughlin[21]将相关性判断分为相关、部分相关和不相关，研究用户在判断三种相关性时所使用的标准以及标准在判断过程中发挥的积极或消极作用。

A. R. Taylor[22－23]将重点放在了随着任务阶段的深入相关性标准使用的变化情况，并在后续的研究中对实验进行了改进，通过技术手段设计了集检索文献与问卷调查为一体的检索系统，用户可以根据每一条检索结果及时作出相关性判断，以及判断所使用的相关性标准。研究结果充分展示了随着 ISP(Information Search Process)阶段的深入相关性标准使用发生的变化。成颖[24]借鉴 A. R. Taylor 博士论文的实验设计，也开展了面向性别差异与任务复杂性的相关性标准研究。研究结果表明，任务复杂性对相关性标准的选择存在显著影响，性别仅仅对文献总体、作者和文献类型3类判据的选择存在影响，总体而言影响甚微。

R. A. Hamid 等[25]研究了不同任务下图像相关性标准对相关性判断的影响。研究结果显示不同任务下，用户使用的相关性标准不同，跨越3个任务都被使

用的标准有 3 个,分别是主题性、信息的吸引度和构成。R. A. Hamid 等[26]对图像相关性标准进行了再次深入研究,将实验任务进一步细化,并深入探讨了熟悉度、困难和满意度对相关性判断的影响。实验选取 48 名用户,综合运用情境检索、调查问卷和录屏的方式记录用户检索行为。实验结果显示在不同实验任务中,用户所使用的相关性标准是相同的,但是相关性标准的重要性却有所不同,如主题性、准确性、构图、情感等。

Y. S. Chang[27]将研究重点放在了在线新闻检索结果网页列表的相关性标准使用的动态性方面。该动态性主要体现在两个方面:一是检索结果网页页数后延相关性标准的使用不同;二是固定网页检索结果排列顺序不同,相关性标准的使用不同。59% 的用户仅使用了单一标准进行了相关性判断,且主题性和来源是两个最主要的标准。

2.3 方法综述

相关性标准实验设计从最初的单一式定性分析发展到现在的多角度定量与定性相结合。研究方法从情境实验、访谈、出声思考、调查问卷等传统方法发展到以情境实验、访谈、出声思考、调查问卷、眼动记录、编程技术抽取、嵌入式系统、日志、民族志等方法相结合使用的三角测量法;分析方法也从定性的内容分析和频次统计分析发展到定量的结构方程、回归分析、相关性分析等。

研究方法的不断完善也使相关性标准的研究更加具有说服力。问卷调查与结构方程相结合可以定量准确分析各个相关性标准对相关性判断的影响;检索日志的获取与分析可以最大化贴近用户的现实检索情境;眼动仪方法的引入可以直观观测用户的注视行为和检索行为。即便如此,依然存在一些问题:①大量数据的定量分析(如调查问卷和检索日志)无法对标具体用户,更脱离了检索情境;②情境实验与访谈耗时较长,定性分析主观性较强。

为了弥补当前研究方法的不足,研究决定引入判断决策领域中的层次分析法,将情境实验与层次分析法相结合,结合情境将每个用户的差异性考虑进行,通过定性打分与定量线性计算,得出相关性判断中相关性标准的权重,使结果更具有说服力。

3 研究思路与实验设计

3.1 研究思路

科学数据用户相关性标准使用特征是用户检索数据过程中,对外界客观数据信息进行认知加工从而产生的外在行为表现。根据相关性标准的研究,认为存在跨越不同情境的共性相关性标准,但是相关性标准的权重和使用频次会随着情境的变化而发生变化。用户使用相关性标准是为了做出相关性判断,因此相关性标准的使用特征放到相关性判断的过程和情境中去测量。相关性标准的重要性是探究用户在长时记忆中较为固定的判断规则,通过调查问卷获取。

在前期针对科学数据相关性标准的研究中,设计了情境实验,对访谈内容进行了内容分析,并根据扎根理论提取了 9 个相关性标准,分别是主题性、便利性、全面性、权威性、质量、规范性、可获得性、可用性、时效性[28]。通过调查问卷将相关性标准进行聚类分析,将 9 个相关性标准分为五大类,其中便利性、可用性和可获得性归为获得成本类,质量、规范性和全面性归为数据优劣类[29]。

本研究在前期相关性标准研究的基础上,设计实验,探究其相关性标准的使用特征。以传统相关性标准的情境研究设计为主线,并引入了判断决策领域的研究方法。结合实际工作任务,采用深度访谈、调查问卷与层次分析法定性与定量相互结合的方法展开科学数据用户相关性标准使用特征研究。

3.2 可操作化

本文研究的 3 个子问题涉及以下几个变量:在相关性判断过程中,用户相关性标准的使用频率、顺序、重要性、权重以及不同情境中相关性标准的权重变化。不同的变量数据要通过不同的研究方法获取。

相关性标准的重要性是用户在长时记忆中对各个标准的认知,通过调查问卷对各个相关性标准进行打分获取,见图 1。

相关性标准的使用频率、顺序、权重以及权重的变化都是在特定的情境下,因此根据用户手头任务开展深度访谈获取相关性标准的使用频率和顺序。访谈结束之后让用户依据自己选择数据的标准用层次分析法进行打分,获取相关性标准的权重以及权重的变化。

其中深度访谈提纲问题有:

(1)个人基本信息:年龄、研究方向、研究成果、检索数据所用平台和频次。

(2)完成当前任务中,需要的数据有哪些?

(3)检索数据时,首先判断的依据是什么? 其次判断的依据是什么?

层次分析法的结构模型见图 2。

*21. 下面的表格中，我们列出了相关性标准的解释，请根据这些解释回忆在检索过程中，进行相关性判断时使用的相关性判断标准有哪些？并根据不同标准对您进行判断的影响程度进行打分。（依据重要程度，从不关注打0分，不重要打1分，不太重要打2分，一般打3分，重要打4分，非常重要打5分）

主题性	0	1	2	3	4	5	查找数据与自己的研究主题相符
可获得性	0	1	2	3	4	5	外界条件的限制不影响自己通过各种渠道获得数据
时效性	0	1	2	3	4	5	在时间上满足自己的研究需要
质量	0	1	2	3	4	5	数据的质量
权威性	0	1	2	3	4	5	在领域里面的认可程度
规范性	0	1	2	3	4	5	符合相应的标准要求
可用性	0	1	2	3	4	5	数据可直接被使用
便利性	0	1	2	3	4	5	方便获取和使用
全面性	0	1	2	3	4	5	数据对研究主题的覆盖程度高

图1　问卷相关性标准重要性打分截图

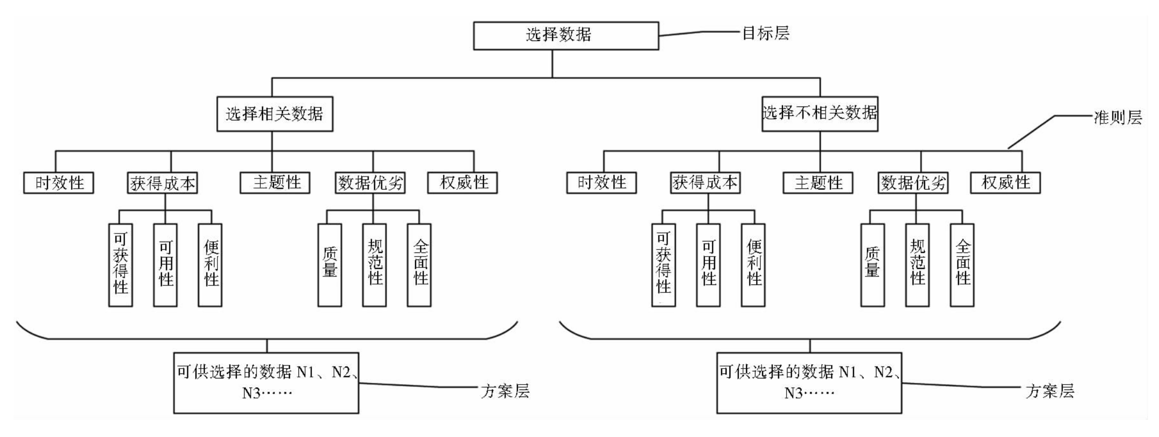


图2　层次分析法结构模型

3.3　实验设计

本实验采用了情境访谈、调查问卷与层次分析法相结合的方法。首先，针对参加科技部举办科技资源共享服务大赛的硕士、博士研究生在线发放问卷，问卷的主要内容包括用户的基本信息、平时检索数据的频率、对科学数据的了解以及用户在完成大赛过程中选取科学数据时相关性标准所起的作用（根据重要性以5级量表的形式打分）；其次，从参与比赛的群体中选取中国农业科学研究院6个研究所的硕士、博士研究生20人，开展访谈，深度调查用户在大赛过程中，选取与自己需求/研究相关数据时所使用的科学数据相关性标准，并追问其对相关性标准的理解以及使用的原因，将访谈结果记录在事先准备好的卡片上。最后，访谈结束后，根据层次分析法对相关性标准进行专家打分，并根据打分计算相关性标准在相关性判断中的权重。

具体来说，研究分为两个阶段：

（1）第一阶段为调查问卷阶段，通过与科技资源共享服务大赛主办方联系，在其QQ群和官方网站发布实验通知和相关问卷填写要求并附带问卷链接，要求硕士及以上学历，且在大赛项目完成过程中在各大数据中心进行了数据检索。对于完成问卷的被试，每人发放10元的劳务费用。

（2）第二阶段为情境访谈实验阶段。从参与共享杯大赛的名单中筛选出中国农业科学院研究生，其研究方向和领域以数据检索为主，最终确定被试20人，分别来自农业资源与农业区划研究所、农业经济与发展研究所、农业信息研究所、生物技术研究所、饲料研究所和农业环境与可持续发展研究所，专业涉及农业经济学、遥感、气象、生态、生物基因、食物营养专业。1－10号为数据一般型用户，研究以实验室数据为主，数据检索目的为验证支持结论，检索数据频率一般。11－20号为数据密集型用户，研究以公开数据为主，数据检索目的是获取研究对象，检索频率密集。前来参加实验的被试，按100元/小时支付费用。

4　数据处理

4.1　问卷数据处理

实验共计收到671份问卷，剩余有效问卷544份。

主要使用 Excel 对相关性标准重要性分值进行算数平均(不考虑个人权重),得到总体群体相关性标准重要性的平均分数。

4.2 访谈数据处理

通过 Nvivo 进行内容分析,根据已经确定的相关性标准和定义对访谈内容进行编码,由课题组两名工作人员进行交叉编码,并保证一致率在 85% 以上。编码结束后统计各个相关性标准在访谈中出现的先后顺序,归纳相关性判断中相关性标准的使用流程,并统计相关性标准出现的频次以及在判断过程中所起的作用,如表 1 所示:

表 1 访谈数据相关性标准使用统计
(以一次相关性判断为例)

标准	是否使用	出现次序	影响程度(分值 1-5)	正向影响	负向影响	相关性
主题性	是	1	5	√		相关
质量	是	2	3		√	不相关
权威性	是	3	1		√	不相关

4.3 层次分析法数据处理

专家打分数据分析主要分为 3 个步骤:

(1)计算每个矩阵各个相关性标准的权重,并保证结果符合一致性检验。

(2)对专家权重进行计算,根据矩阵的特征向量确定专家的内容权重;根据判断矩阵的一致性比例确定专家的逻辑权重;两者求几何平均值得到专家最终的权重[30]。

根据结构模型可以看出,用户构建了 6 个打分矩阵(相关数据 3 个,不相关数据 3 个)。因此分别得到表 2、表 3 所示的专家权重。

(3)专家权重与相应相关性标准权重相乘求和,最终得到每个相关性标准的权重,见表 4。

5 结果分析

通过对两个阶段实验数据的分析比较,可以发现科学数据用户相关性标准使用的 4 个特征。

表 2 层次分析法 1-10 号用户权重

相关性	矩阵类型	1 号	2 号	3 号	4 号	5 号	6 号	7 号	8 号	9 号	10 号
相关数据	获得成本层	0.077 2	0.070 1	0.060 5	0.075 8	0.039 5	0.158 8	0.099 9	0.110 2	0.158 6	0.149 4
	数据优劣层	0.076 6	0.029 6	0.072 5	0.137 1	0.024 2	0.151 3	0.150 1	0.135 2	0.132 5	0.090 8
	选择数据层	0.122 0	0.113 6	0.090 9	0.042 1	0.053 6	0.136 7	0.069 6	0.108 8	0.126 3	0.136 4
不相关数据	获得成本层	0.172 1	0.117 0	0.098 6	0.112 2	0.043 2	0.166 7	0.039 2	0.036 4	0.137 3	0.077 4
	数据优劣层	0.040 8	0.049 7	0.131 7	0.131 7	0.142 4	0.145 4	0.112 9	0.149 2	0.047 8	0.048 4
	选择数据层	0.143 2	0.176 6	0.039 4	0.032 0	0.046 3	0.160 4	0.107 8	0.121 1	0.119 5	0.053 8

表 3 层次分析法 11-20 号用户权重

相关性	矩阵类型	11 号	12 号	13 号	14 号	15 号	16 号	17 号	18 号	19 号	20 号
相关数据	获得成本层	0.083 5	0.084 1	0.120 4	0.080 7	0.080 0	0.0802	0.1230	0.113 6	0.123 3	0.111 4
	数据优劣层	0.153 8	0.076 9	0.153 8	0.076 9	0.153 8	0.076 9	0.153 8	0.076 9	0.038 5	0.038 5
	选择数据层	0.089 7	0.157 5	0.045 9	0.157 7	0.105 0	0.123 7	0.108 3	0.108 9	0.048 7	0.054 7
不相关数据	获得成本层	0.077 1	0.121 2	0.079 7	0.121 2	0.118 0	0.104 6	0.119 9	0.072 7	0.078 5	0.107 0
	数据优劣层	0.020 4	0.114 9	0.153 3	0.153 3	0.111 8	0.019 0	0.096 9	0.149 3	0.109 3	0.071 9
	选择数据层	0.035 2	0.180 7	0.152 4	0.061 7	0.129 7	0.109 5	0.101 8	0.098 4	0.099 0	0.031 5

表 4 不同情境下相关性标准权重

相关性判断	相关数据		不相关数据	
	数据一般型	数据密集型	数据一般型	数据密集型
时效性	0.143 5	0.125 9	0.139 2	0.116 3
可获得性	0.042 2	0.038 3	0.004 7	0.038 7
可用性	0.064 2	0.046 2	0.060 8	0.052 4
便利性	0.026 1	0.018 5	0.032 6	0.019 1
主题性	0.264 5	0.345 1	0.315 8	0.349 0
质量	0.124 6	0.112 7	0.123 9	0.098 3
规范性	0.040 4	0.057 8	0.053 2	0.049 3
全面性	0.078 7	0.067 0	0.062 4	0.074 3
权威性	0.193 8	0.188 5	0.175 5	0.202 6

5.1 相关性标准的使用频次呈现长尾态势且满足最少省力原则

通过对用户的访谈内容进行分析,计算各个相关性标准出现的频次,并绘制了百分比分布图,如图 3 所示。由分布图可以看出,使用最为频繁的标准为主题性(36.8%),其次为质量(23%),排名第三位的为可获得性(11%),主题性、质量和可获得性的使用频率占据 2/3,这 3 个标准被频繁使用。而剩下 6 个相关性标准每个标准的占比不超过 10%,甚至可用性、权威性和规范性的占比均低于 5%,可见这几个标准偶尔

被使用。但是6个相关性标准出现频率的总和却达到了30%，占据了将近1/3的比例。整体呈现出长尾分布的态势。

结合调查问卷和情境实验，对比分析了相关性判断中，相关性标准的使用频次和重要性分数，结果显示两者并非正相关，频次分布图与重要性分布图趋势差异较大，如图3所示。权威性在重要性打分中分值为4.06，排在第三位，仅次于质量和主题性，但是在实际情境检索过程中，使用频次仅为7次，排在倒数第二位。可获得性在重要性打分中，分值为3.71，排在倒数第二位，但是在实际情境检索过程中，使用频次为33次，排在第三位。同样规范性、便利性等也存在同样的现象。结合访谈内容得出，用户在实际检索过程中，更希望以最少的时间和精力投入获取理想的结果。如果最优的数据(高质量、高权威性)需要投入成本过大，则会降低要求，选择替代数据。

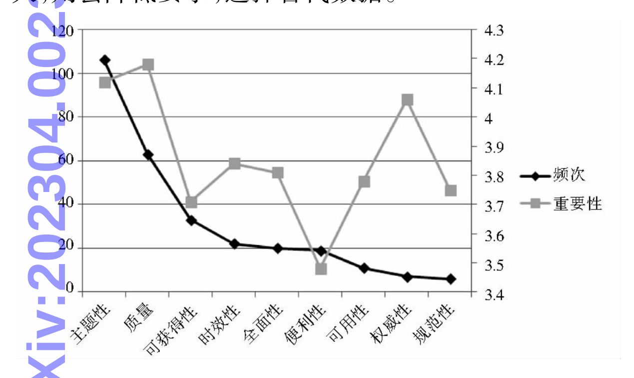


图3 相关性标准使用频次和重要性对比分布

5.2 主题性在相关性判断过程中具有锚定功能

在相关性判断中，最常用的相关性标准为主题性、质量和可获得性。这些标准不是同时出现的，而是有先后顺序的。在一般情况下，用户最先是根据关键词、题目和摘要等线索判断科学数据的主题性。在主题性相关的前提下，再根据其他标准进行调整。

如表5所示，访谈中主题性出现总频次106次，以第一次序的形式出现的频次为106次。说明在一次判断中，只要使用了主题性标准进行判断，那肯定是第一位使用的标准。出现次序为1的标准还有可获得性和便利性。追溯访谈内容发现，可获得性出现在第一位的判断情况是检索结果均有获取权限。便利性出现在第一位的判断情况是检索结果需要用户注册才能查看，用户直接放弃。由此可见在数据资源充足可用的前提下，用户进行数据相关性判断的首要使用标准是主题性。在一次判断中，用户一般使用3－5个标准进行判断，从第二个使用的标准开始，就会依据个人喜好、认知情况、任务情况而有所区别。主题性确定之后，主题性相关，则进一步使用质量、可获得性等标准判断；若主题性不相关，则直接放弃，开始下一轮检索判断。这是主题性的使用频次远高于其他标准的原因。因为主题性不仅用来判断相关数据，更用来否定不相关数据。

表5 相关性标准出现次序频次统计 （单位：次）

相关性标准	出现总频次	出现次序为1	出现次序为2	出现次序为3
主题性	106	106	0	0
质量	63	0	34	21
可获得性	33	7	15	7
时效性	22	0	2	8
全面性	20	0	0	11
便利性	19	1	0	1
可用性	11	0	2	7
权威性	7	0	3	2
规范性	6	0	0	0

因此这符合判断决策中的锚定与调整策略，用户在判断开始阶段将注意力集中在主题相关这个锚上，在后续的判断中，根据其他的标准对这个锚进行不充分调整。若锚不成立，则需要重新进行锚定或直接结束判断得出“不相关”的结果，如图4所示[31]：

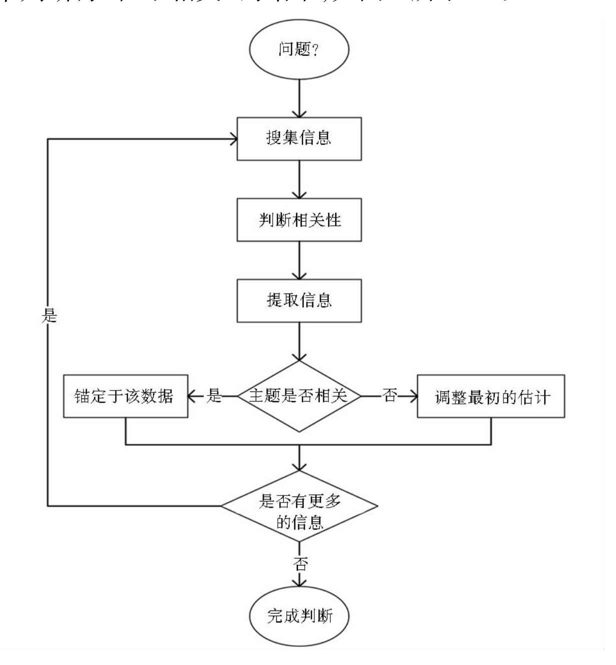


图4 科学数据相关性判断中相关性标准锚定与调整流程

当用户开展一项数据任务时，根据自己的检索任务，会首先在心中有个预期，并设定好检索式。在检索过程中，根据检索式检索出来的结果与自己内心的预期结果进行比对来判断相关性。用户总会首先关注关

键词、题目、摘要等一系列与主题相关判断有关的信息。若主题相关,则锚定该条信息,并根据后续更多的信息判断质量、可获得性、权威性、可用性等,并在主题相关的基础上,对判断结果进行调整。若主题不相关,则需要调整最初的估计,开始新一轮检索。

5.3 相关性标准具有双向性

在相关性判断中,相关性标准的作用具有双向性。早在 2002 年 K. Maglaughlin 学者的研究中就提到相关性标准对文献相关性判断不仅有积极影响,还有消极影响[21]。同样,针对科学数据相关性判断也有相同的特征。科学数据相关性标准对相关性判断的影响可以分为两类,一类是主题性、时效性、权威性、质量、全面性和规范性,均具有双向性。如时效性,用户说"这个是发布时间刚好是我想要的",就起到了积极作用;用户说"这个时间太老了,已经不能用了",就起到了消极作用。另一类是可获得性、可用性和便利性,均为消极影响,只不过影响程度有所差异。如可获得性,用户说"这是数据不能下载,无法进一步判断",就起到了消极作用。探究其原因,主要是因为第一类为衡量科学数据元数据方面的标准,第二类为衡量科学数据共享系统方面的标准。在对共享系统进行判断时,只有不太顺利才会引起用户的注意,当检索数据比较顺利的时候,用户会忽略这些方面。

5.4 相关性标准的使用权重相对稳定

不同类型用户在判断数据相关与不相关时,虽然相关性标准权重数值存在差异,但是整体使用权重差异不显著。依据专家打分和层次分析法计算,最终得到数据一般型用户和数据密集型用户在判断数据相关以及数据不相关时各个相关性标准的权重,见表 4。

不同类型用户在相同的数据判断结果中具体标准权重分配上存在差异,如数据一般型用户选择相关数据时,主题性权重为0.264 5,权威性权重为 0.193 8;数据密集型用户选择相关数据时,主题性权重为 0.345 1,权威性权重为 0.188 5。相同类型用户在不同的数据判断结果中具体标准权重分配上存在差异,如数据一般型用户在判断数据相关时,主题性权重 0.264 5,权威性权重 0.193 8;在判断数据不相关时,主题性权重 0.315 8,权威性权重 0.175 5。由此可见,判断数据相关时,主要判断标准(如主题性、质量、权威性和时效性)的权重分配较为均匀。而判断数据不相关时,主要判断标准的权重分配差距明显拉大。主题性权重占据了 41.6%,也体现出在判断不相关时,主题性的作用更大。但是无论是不同用户类型选择相关数据的权重对比还是相同用户类型选择不同数据的权重对比,整体上相关性标准发挥作用的整体趋势一致,且通过配对 T 检验,发现显著性系数 P 均大于 0.01,接受原假设(见表 6),即相关性标准权重整体没有显著差异。

表 6 不同对组间相关性标准权重配对 T 检验

对组	一般型用户相关 - 不相关	密集型用户相关 - 不相关	相关—一般型 - 密集型	不相关—一般型 - 密集型
显著性	0.867	0.997	0.888	0.999

因此可以看出相关性标准权重具体数值随着情境的变化会有差异,标准的使用顺序也会有所不同,但是其整体使用权重会保持相对稳定。

6 讨论

本项研究立足科学数据开展研究,探究了在科学数据相关性判断过程中,相关性标准的使用特征。研究采用了决策领域的层次分析法与本领域研究传统的情境实验和调查问卷相结合的方式开展研究。情境实验是相关性研究领域中最为经典的实验设计(C. Watson[32]、A Hamid 等[25]、高飞[16]),结合情境开展实验既可以充分将用户需求与情境相结合,更加真实地反映判断过程,又可以即时获取想要的数据,保证数据的质量。这种方法获取的数据主要依靠内容分析进行定性分析,学者们主要统计了相关性标准出现的频次,依据频次来阐述其重要性,但是本文通过数据分析发现,相关性标准的使用频次与重要性不成正比。因此此种方法不能很好地探究其对相关性判断的影响。为了更加客观地分析相关性标准的影响,有学者(Z. Chen 等[33]、Y. Xu[5]、韦草原等[34])引入了结构方程,通过大量的问卷调查数据建立结构方程模型,从而得到相关性标准的权重,解决了情境实验面临的困境。但是相关性判断既与情境密切相关,又具有个性化差异,脱离情境的问卷调查会丧失数据的真实性,也无法体现用户之间的差异。考虑到这些问题,本研究引入了决策领域的层次分析法,既能贴合任务情境和用户的个性化差异,又能进行定量计算,分析相关性标准的权重。从而弥补了传统方法的不足。

用户在判断科学数据过程中,相关性标准的使用情况与传统信息载体基本一致。在文献、图像、网页、视频等一系列的研究中,主题性和质量都是重要的标准[35-36]。但是也存在一些差异,比如时效性的影响就比较低,因为在很多研究者眼中,数据没有过时,重点在于自己关注的时间范围,有时候甚至年代越久远的

数据对于研究更有价值。这一点与文献、新闻等就有很大差异。同时也体现了数据管理与存储的重要性。

同时相关性标准使用符合长尾定律。当前数据共享如火如荼，但是却存在诸多壁垒。虽然用户常用的相关性标准只有3个，但是剩下的标准却占据了30%的份额。很多数据共享的困境正是由这些小份额的相关性标准造成的，如数据不规范、下载的数据不全、数据格式不统一等。如果数据共享与服务的研究中可以充分考虑这些小占比的相关性标准，破除壁垒，将极大提升数据共享与服务质量。

锚定调整策略广泛应用于判断决策领域，如司法判断、财务问题等。相关性判断是信息科学领域的基本过程，与情境有着密切联系，同样满足锚定调整策略。主题性为锚，也是相关性判断中最重要的标准，最终的判断结果是根据情境需要不断在主题性这个锚定的基础上不断调整得出的。

通过层次分析法分别计算了标准在判断数据相关与数据不相关时的权重。Y. Xu[5]用结构方程计算了5个文献相关性标准（主题性、可信度、可理解性、新颖性、领域）的权重，但是没有对判断结果进行分类计算。其中主题性为0.31，与本文主题性的权重相近。S. Y. Rieh[37]通过定性研究得出除了主题性，质量和权威性是更为重要标准的结论。可以看出，通过层次分析法计算得出的权重基本与其他学者研究中相关性标准的权重分布大致一致。因此可以看出相关性标准的使用权重相对保持稳定，也进一步证实了将层次分析法引入到相关性判断领域研究的可行性和可操作性。

7 结论

本文在前人研究提出的9个科学数据相关性标准的基础上开展研究，通过调查问卷、深度访谈和层次分析相结合的方法探究了科学数据相关性标准的使用特征。研究表明，科学数据相关性标准使用主要呈现以下4个特征：①相关性标准的使用情况符合著名的长尾定律且满足最少省力原则；②主题性在相关性判断过程中具有锚定功能；③相关性标准在相关性判断中具有双向性；④相关性标准使用权重相对稳定。

通过对科学数据相关性标准使用特征的探索，准确把握了科学数据用户的检索策略，明晰了其检索行为，从而为改进科学数据共享系统提供可操作性的建议如下：

（1）补充完善科学数据共享平台中的缺失信息。用户反馈在国内数据共享平台检索时，很多数据都不够全面，一些数据只有简介，具体数据都是空白的，为他们后续的数据使用造成了很多麻烦。因此需要提高我国科技人员数据共享的意识，也不断加强与完善数据共享基础设施，补充完善缺失数据，切实为用户提供数据支撑。

（2）完善平台的“导航式”检索模式。用户在检索数据中以主题判断为主，访谈过程中也发现用户更愿意使用导航检索。导航式检索可以帮助用户更快地定位到相关的领域和主题。因此要针对各大平台的数据特点和用户人群，对数据进行精确分类，从而辅助用户快速检索。

（3）改进科学数据共享平台中元数据的页面布局。科学数据包含很多元数据，不同元数据在用户判断过程中起的作用不同，与主题相关的题目、摘要、关键词，与质量相关的数据采集过程、采集标准、认证级别，与权威性相关性的发布单位、发布者等元数据更为重要，应放在醒目的位置，方便用户判断。

（4）改进系统的人机交互性。根据IP或者用户名记录每位用户的检索记录，根据浏览和点击记录不断明确用户的检索策略，根据不同的检索策略为不同的用户提供不同的信息呈现顺序和方式，从而满足个性化需求。

参考文献：

[1] CUADRA C A, KATTER R V. Opening the black box of relevance [J]. Journal of documentation, 1967,23(2):291－303.
[2] REE A M, SCHULTZ D. G. A field experimental approach to the study of relevance assessments in relation to document searching [R]. Cleveland: Case Western Reserve University,1967.
[3] BARRY C, SCHAMBER L. Users' criteria for relevance evaluation: a cross-situational comparison[J]. Information processing & management,1998,34(2/3):219－236.
[4] TAYLOR A R, COOL C, BELKIN N J, et al. Relationships between categories of relevance criteria and stage in task completion [J]. Information processing & management, 2007, 43(4): 1071－1084.
[5] XU Y, CHEN Z. Relevance judgment: what do information users consider beyond topicality? [J]. Journal of the Association for Information Science & Technology, 2010, 57(7):961－973.
[6] BARRY C. User-defined relevance criteria: an exploratory study [J]. Journal of the American Society for Information Science, 1994, 45(3): 149－159.
[7] SCHAMBER L, EISENBERG M B, NILAN M S. A re-examination of relevance: toward a dynamic, situational definition[J]. Information processing & management,1990,26(6):755－776.

[8] COOL C, BELKIN N, FRIEDER O, et al. Characteristics of text affecting relevance judgments[C]//WILLIAMS M E. Proceedings of the 14th National Online Meeting, New York: Learned Information Inc., 1993: 77-81.
[9] WANG P, SOERGEL D. A cognitive model of document use during a research project. Study I. Document selection[J]. Journal of the American Society for Information Science, 1998, 49(2): 115-133.
[10] WANG P, WHITE M D. A cognitive model of document use during a research project. study II. decisions at the reading and citing stages[J]. Journal of the Association for Information Science & Technology. 1999,50(2):98-114.
[11] ALONSO O, MIZZARO S. Relevance criteria for e-commerce: a crowdsourcing-based experimental analysis[C]//Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval. Boston: ACM, 2009: 760-761.
[12] SABBATA D S. Criteria of geographic relevance[C]// WEIBEL R. Fabrikant S. GIScience 2010 - 6th International conference on geographic information science. Zurich: Springer, 2010: 1-4.
[13] SABBATA D S, REICHENBACHER T. Criteria of geographic relevance: an experimental study[J]. International journal of geographical information science, 2012, 26(8): 1495-1520.
[14] REICHENBACHER T, SABBATA D S, PURVES R S, et al. Assessing geographic relevance for mobile search: a computational model and its validation via crowdsourcing[J]. Journal of the Association for Information Science and Technology, 2016, 67(11): 2620-2634.
[15] WANG X, HONG Z, XU Y C, et al. Relevance judgments of mobile commercial information[J]. Journal of the Association for Information Science and Technology, 2014, 65(7):1335-1348.
[16] 高飞. 科学数据用户相关性线索、标准及相互关系研究[D]. 北京:中国农业科学院,2017.
[17] 张贵兰, 王健, 刘建平, 等. 相关性判据与目标信息类型的关系分析[J]. 情报杂志, 2018,37(6):171-179.
[18] BATEMAN J. Changes in relevance criteria: a longitudinal study [C]//Proceedings of the ASIS annual meeting. Pittsburgh: ERIC, 1998: 23-32.
[19] BATEMAN J. Modeling the importance of end-user relevance criteria[C]//Proceedings of the ASIS annual meeting. Washington: ERIC, 1999: 396-406.
[20] VAKKARI P, HAKALA N. Changes in relevance criteria and problem stages in task performance[J]. Journal of documentation, 2000, 56(5): 540-562.
[21] MAGLAUGHLIN K L, SONNENWALD D H. User perspectives on relevance criteria: a comparison among relevant, partially relevant, and not-relevant judgments[J]. Journal of the American Society for Information Science and Technology, 2002, 53(5): 327-342.
[22] TAYLOR A, ZHANG X, AMADIO W J. Examination of relevance criteria choices and the information search process[J]. Journal of documentation, 2009, 65(5): 719-744.
[23] TAYLOR A. User relevance criteria choices and the information search process[J]. Information processing & management, 2012, 48(1):136-153.
[24] 成颖. 面向性别差异与任务复杂性的相关性判据研究[J]. 图书情报工作,2011,55(18):29-35.
[25] HAMID R A, THOM J A. Criteria that have an effect on users while making image relevance judgements[C]//Fifteenth Australasian document computing symposium, Melbourne: RMIT University, 2010: 1-8.
[26] HAMID R A, THOM J A, ISKANDAR D N F A. Effects of relevance criteria and subjective factors on Web image searching behaviour[J]. Journal of information science, 2017, 43(6): 786-800.
[27] CHANG Y S, GWIZDKA J. Relevance criteria dynamics: a study of online news selection on SERPs[J]. Proceedings of the Association for Information Science and Technology, 2018, 55(1): 768-769.
[28] 张贵兰,王健,周国民,等. 科学数据用户相关性标准研究[J]. 图书情报工作,2019,63(4):112-121.
[29] ZHANG G L, WANG J, ZHOU G M, et al. Scientific data relevance criteria classification and usage[C]//Proceedings of 2nd international conference on computer science and application engineering, Huhhot:ACM,2018:1-7.
[30] 蒋文能. 群组决策中专家权重确定的思路和方法[J]. 统计与决策, 2013 (2): 24-28.
[31] 张贵兰. 科学数据用户相关性标准研究[D]. 北京:中国农业科学院,2019.
[32] WATSON C. An exploratory study of secondary students' judgments of the relevance and reliability of information[J]. Journal of the Association for Information Science and Technology, 2014, 65(7): 1385-1408.
[33] CHEN Z, XU Y. User-oriented relevance judgment: a conceptual model[C]//Proceedings of the 38th annual Hawaii international conference on system sciences. Big Island:IEEE, 2005: 1-11.
[34] 韦草原, 王健, 张贵兰, 等. 基于扎根理论的科学数据用户感知价值概念模型研究[J]. 情报杂志, 2018, 37(5):186-192.
[35] BALATSOUKAS P, RUTHVEN I. An eye-tracking approach to the analysis of relevance judgments on the Web: the case of Google search engine[J]. Journal of the American Society for Information Science and Technology, 2012, 63(9): 1728-1746.
[36] KIM Y W, PARK S J. User-based relevance and irrelevance criteria during the task pursuing of middle school students[J]. Journal of the Korean Society for Library and Information Science, 2014, 48(3): 55-70.
[37] RIEH S Y. Judgment of information quality and cognitive authority in the Web[J]. Journal of the American Society for Information Science and Technology, 2002, 53(2): 145-161.

张贵兰，周国民，潘尧，等.科学数据相关性标准使用特征[J].图书情报工作,2020,64(10):56－65.

作者贡献说明：

张贵兰：整体论文框架确定,数据分析,论文撰写与修改；
周国民：论文修改；
潘尧：数据采集与分析；
刘建平：整理相关资料,协助数据采集；
王健：研究思路提出,论文修改意见提出。

Usage Characteristics of Scientific Data Relevance Criteria

Zhang Guilan[1] Zhou Guomin[2] Pan Yao[2] Liu Jianping[2] Wang Jian[2]
[1] Institute of Scientific and Technical Information of China, Beijing 100038
[2] Agricultural Information Institute of CAAS, Beijing 100081

Abstract: [**Purpose/significance**] This paper explored the use characteristics of relevance criteria that affect the users' relevance judgment when they retrieved data, so as to further explore the information behavior of scientific users and provide theoretical support for improving the data retrieval platform. [**Method/process**] The research was carried out by combining questionnaire, interview and AHP. The subjects were graduate students in universities and research institutes. [**Result/conclusion**] The analysis results showed that the relevance criteria presented 4 characteristics: the use of relevance criteria conformed to the famous long tail law; topicality had anchoring function in the process of relevance judgment; relevance criteria were bidirectional; the weights of relevance criteria were relative stability. The research accurately grasped the retrieval strategy of scientific data users and clarified their retrieval behavior, thus providing operable suggestions for improving the scientific data sharing system.

Keywords: scientific data　relevance criteria　use characteristics of relevance criteria　user behavior